BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

III. Examiner Interview Summary

In a telephone conversation of April 19, 2004 between Examiner Sisson and Gloria L. Norberg, Examiner Sisson clarified the due date of a response to the Advisory Action as allowing a further one month extension of time to May 4, 2004.

Examiner Sisson also clarified a sentence of the Advisory Action as needing the word "are" in place of "re" in the sentence: The Action notes that the plurality of "target receptors of length up to 10 microns" re **(are, sic)** also to be "single-stranded nucleic acids of a predetermined base sequence ...[and] have the same or different base sequence.

In the telephone conversation, Examiner Sisson also stated that the rejection applies to double stranded embodiments of the invention. The examiner is looking for evidence that the invention as claimed was in the possession of Applicants. The examiner believes that no such sequences or characteristics of such sequences are set forth in the specification.

The examiner stated that the written description may be found adequate for process claims.

Further, the University of Rochester case was cited by the examiner as setting forth written description requirements. University of Rochester v. G.D. Searle & Co., 69 USPQ2d 1886 (CA FC 2004).

Applicants' representative appreciates the courtesy of the examiner in returning her telephone call and clarifying the rejection.

IV. Remarks

A. Status of the Application

Claims 35-37 are amended. Claims 8-11 and 35-37 are pending.

B. Rejections of Claims 8-11 and 35-37 under 35 U.S.C. §112, First Paragraph

Office Action

Claims 8-11 and 35-37 were rejected in the Final Office Action for introduction of new matter. In particular, Claims 35-37 were rejected for the following: "where target receptors have a length of up to 1mm," "a labeled complex having a predetermined molar ratio of the labeled substances," and "number and length of target receptors."

Claim 35 was rejected in the Final Office Action for "where single-stranded target receptors have a predetermined base sequence."

The Advisory Action states that the response filed December 29, 2003 does not place the application in condition for allowance because: a review of page 9 and of original claim 3 fails to find support for where the "target receptors," which "are single-stranded nucleic acids of a predetermined base

sequence are "of length up to 1mm." Support is cited as present for where the combination of "target receptor, which is bonded with the carrier on a part thereof, and bonded with the labeled substance on the other part thereof, is formed in a slender shape" which can be of a length of up to 1mm. The combination of elements was cited as not teaching that the "target receptor" alone, which must be of a predetermined sequence and be single stranded nucleic acid, be of that length.

The Communication mailed February 27, 2004 stated that removal of the "length up to 1mm" language would remove the new matter rejection but would raise a new issue of total reliance on functional language as it relates to satisfaction of the written description requirement.

The Advisory Action mailed April 14, 2004 states that the proposed limitation of 10 microns on the length of the target receptors is not adequately supported by the specification in terms of written description. The Action notes that the plurality of "target receptors of length up to 10 microns" re (are, sic) also to be "single-stranded nucleic acids of a predetermined base sequence ...[and] have the same or different base sequence.

The Advisory Action mailed April 14, 2004 also cites U.S. Patent 6,465,241 to Haronian *et al.* as teaching that the length of axial rise per nucleotide in DNA is 3.3 Angstroms, or 3.3 x 10-4 micrometers and, in view of this teaching, Applicant's single stranded nucleic acid of a predetermined length would be 30,303 nucleotides long. A review of the disclosure, including applicant's remarks is cited as failing to find an adequate written description of multiple target receptors of this length.

Response

The responses filed December 29, 2003, January 30, 2004 and April 1, 2004 are requested to be entered into the record of this application in addition to the present response.

Independent Claim 35 has been amended to recite "wherein said target receptors are single-stranded nucleic acids of predetermined base sequence, wherein the single–stranded nucleic acid is a base sequence of a gene, a base sequence of mRNA, a base sequence of tRNA, a base sequence of rRNA, a base sequence obtained by denaturation of a double stranded nucleic acid, or a base sequence obtained by synthesis."

Independent Claims 36 and 37 have been amended to recite "wherein the double–stranded nucleic acid is a base sequence of a gene, a base sequence of mRNA, a base sequence of tRNA, a base sequence of rRNA, a base sequence obtained by using the polymerase chain reaction, a base sequence having a recognition sequence of a restriction enzyme at one end, a base sequence generated by annealing, or a base sequence generated by DNA ligase."

Support for added language is found in the specification as follows: support for "wherein the single–stranded nucleic acid is a base sequence of a gene" is found at page 7, line 6; for "a base sequence

of mRNA," is at page 10, line 14; for "a base sequence of tRNA" is found at page 10, line 14; for "a base sequence of rRNA" is found at page 10, line 14; for "a base sequence obtained by denaturation of a double stranded nucleic acid" is found at page 11, line 8, page 21, lines 7-8, page 34, lines 13-14, and at page 16, line 17-20; for "a base sequence obtained by synthesis" is found at page 15, lines 15-17; for "wherein the double–stranded nucleic acid is a base sequence of a gene" is found at page 7, line 6; for "a base sequence of mRNA, a base sequence of tRNA, a base sequence of rRNA" is found at page 10, line 14; for "a base sequence obtained by using the polymerase chain reaction" is found at page 27, lines 9-14, page 28, lines 22-28, and page 42, lines 25-28; for "a base sequence having a recognition sequence of a restriction enzyme at one end" is found at page 28, lines 23-24; for "a base sequence generated by annealing" is found at page 15, line 19; and for "a base sequence generated by DNA ligase" is found at page 16, line 11, and at page 29, lines 5-6. Applicants submit that no new matter has been entered by the amendments to the claims.

PTO's *Guidelines for Examination of Patent Applications Under the 35 USC 112, P1, "Written Description" Requirement,* 66 Fed. Reg. 1099

The Guidelines state that an applicant shows possession of the claimed invention by describing the claimed invention with all of its limitations using such descriptive means as words, structures, figures, diagrams, and formulas that fully set forth the claimed invention. Possession may be shown in a variety of ways including description of an actual reduction to practice, or by showing that the invention was 'ready for patenting' such as by the disclosure of drawings or structural chemical formulas that show that the invention was complete, or by describing distinguishing identifying characteristics sufficient to show that the applicant was in possession of the claimed invention. The description need only describe in detail that which is new or not conventional.

Factors to be considered in determining whether there is sufficient evidence of possession include the level of skill and knowledge in the art, partial structure, physical and/or chemical properties, functional characteristics alone or coupled with a known or disclosed correlation between structure and function, and the method of making the claimed invention. Disclosure of any combination of such identifying characteristics that distinguish the claimed invention from other materials and would lead one of skill in the art to the conclusion that the applicant was in possession of the claimed species is sufficient. Patents and printed publications in the art should be relied upon to determine whether an art is mature and what the level of knowledge and skill is in the art.

Briefly, the present invention is to a labeled complex comprising three components:

a carrier particle,

a number of target receptors, and

at least a first type and a second type of labeled substance thereby forming a labeled complex having a predetermined molar ratio of the types of labeled substances.

The written description requirement that an applicant must show possession of the claimed invention by describing the claimed invention with all of its limitations using such descriptive means as words, structures, figures, diagrams, and formulas that fully set forth the claimed invention is met by the descriptive words of the amended claims and support as provided *supra*. Further, the structures and figures of the application depict nucleic acids generically as simple lines because their particular sequence is of no consequence.

The written description requirement that possession may be shown in a variety of ways including description of an actual reduction to practice, or by showing that the invention was 'ready for patenting' such as by the disclosure of drawings or structural chemical formulas that show that the invention was complete, or by describing distinguishing identifying characteristics sufficient to show that the applicant was in possession of the claimed invention and that the description need only describe in detail that which is new or not conventional is met by the support cited in the specification for the amended claim language. Cloning and sequencing of nucleic acids so as to provide a nucleic acid of predetermined sequence of a length up to 10 microns was not new as of the priority date of the application as demonstrated *infra*.

The written description requirement that sufficient evidence of possession include the level of skill and knowledge in the art, partial structure, physical and/or chemical properties, functional characteristics alone or coupled with a known or disclosed correlation between structure and function, and the method of making the claimed invention, and that disclosure of any combination of such identifying characteristics that distinguish the claimed invention from other materials and would lead one of skill in the art to the conclusion that the applicant was in possession of the claimed species is sufficient, is met by the presently claimed invention since the particular sequence of a nucleic acid is not of consequence to the present invention. The invention relates to properties common to nucleic acids in general. For example, linear nucleic acids have 3' and 5' ends, they contain sugars and phosphate bonds. Nucleic acids hybridize to their complementary strand by well known rules. Nucleic acids bind proteins that have binding specificity for a particular sequence of the nucleic acid under conditions readily determined by one of skill in the art. The present invention provides a labeled complex that includes nucleic acids and relies on those general properties of nucleic acids that are common to nucleic acids. A requirement for a specific genus or species of nucleic acid or specific sequences of nucleic acid is simply not applicable to

the present invention. Applicants submit that the amended claims recite nucleic acids as specifically described in the specification. To limit the target receptor single stranded or double stranded nucleic acids of the claimed labeled complexes to a species or genus would be similar to limiting the DNA fragments produced in DNA sequencing operations as patented by Caltech (U.S. Patent 5,171,534, Dec. 15, 1992 (the Applied BioSystems Automated DNA analyzer)) to only those that are listed in the patent which is clearly not the scope of the Caltech invention.

The written description requirement that patents and printed publications in the art demonstrate that the art is mature is met by the following disclosures of nucleic acids having predetermined sequences of up to 30.3 kb:

Attachment 1: 1997 Human Genome Program Report: Research Narratives LANL, Los Alamos National Laboratory. Under "Accomplishments" on page 2, the third bullet point states "Generation of finished sequence for a 240,000-base telomeric region of human chromosome 7q.

Attachment 2: 1997 Human Genome Program Report: Research Narratives LLNL, Lawrence Livermore National Laboratory. Under "Accomplishments" on page 4, the third bullet point states "Complete sequencing of human genomic regions containing two additional DNA repair genes. The second human repair gene, HHR23A, maps to 19p13.2. Sequence analysis of 110,000 bases containing HHR23A identified six other genes, ..."

Attachment 3: U.S. Patent No. 6,022,716, filed April 10, 1998, a date that is prior to the priority date of the present application, teaches methods for sequencing of a 145 kb insert of human genomic DNA in a BAC clone beginning at column 31, line 62-64, and methods for sequencing of a 100 kb insert in a BAC clone beginning at column 44, line 45-47.

Therefore, one of ordinary skill in the art would have been able to determine, as of the priority date of the present application, the sequence of a nucleic acid as long as 30.3 kb or 10 microns and, therefore, possess a nucleic acid of such predetermined sequence. Single stranded nucleic acid can be obtained, for example, simply by denaturing double stranded nucleic acid.

Applicants respectfully request that the rejection of the claims for lack of adequate written description be withdrawn.

Remarks as Provided in the Response Filed April 1, 2004

The claims have been amended to recite the length of the target receptor as up to 10 microns. Support for up to 10 microns for the length of the target receptor is found in the specification at page 8, line 28, to page 9, line 9, specifically at line 9. Said lines state:

..., the target receptor, ..., is formed in a slender shape ... . (page 9, line 3). The size of the "slender shape" is not expressly defined (page 9, line 5). ... For example, the form is as long as or sufficiently longer than the particle size (page 9, lines 7-8), for example, about 10 times as long as the particle size, for example, about 10μm (page 9, lines 8-9).

Support for the remaining objectionable claim language of the independent claims is found in the specification and originally filed claims as follows:

In the third aspect of the invention (begins on page 8, last line), the target receptor is formed in a slender shape (page 9, line 1-3). Further, the target receptor is alone referred to as being in a slender shape since the rest of the sentence is set off in phrases with commas. Therefore, the "slender shape" refers to the receptor only.

In the fifth aspect of the invention (begins on page 10, line 18), which refers to any one of the first through the fourth aspects of the invention, the target receptor has a predetermined double strand base sequence. Therefore, this aspect refers to the third aspect which describes the target receptor in a slender shape. Therefore, the receptor may be a double strand base sequence of predetermined sequence in a slender shape.

In the seventh aspect of the invention (begins on page 11, line 11), which refers to any one of the first through the fifth aspects of the invention, the target receptor is a single strand nucleic acid. Therefore, this aspect refers to the fifth aspect which describes the receptor as double stranded, and the seventh aspect further describes the target receptor as denatured to a single strand. This aspect still has a predetermined base sequence and a slender shape.

Therefore, the seventh aspect of the invention describes a target receptor as having a slender shape, as being a single strand nucleic acid, and having a predetermined base sequence.

With regard to the length of a "slender shape," at page 9 and in Claim 3 as originally filed, the target receptor is cited as formed in a slender shape (page 9, line 3). The size of the "slender shape" is not expressly defined (page 9, line 5), however, for example, the form is as long as or sufficiently longer than the particle size (page 9, lines 7-8), for example, about 10 times as long as the particle size, for example, about 10μm (page 9, lines 8-9).

At page 5, the particle size is cited as preferably of the order of about 0.1μm ~ about 1 mm (lines 11-13). Therefore, the target receptor may be from 0.1 μm to about 10 mm. The Office Action of Feb. 12, 2003 requested that a limit be put on the length of immobilized polynucleotides. Since the value of 10 microns is expressly stated in the specification, that value is selected as a limit on the length of target receptors.

The receptor, in a slender shape, functions as a spacer (page 9, lines 10-11). As supported by page 9, lines 10-23, the "slender shape" and, therefore, the length of the receptor is such that a major

influence by energy movement or quenching among the labeled substances does not occur, thereby enhancing discrimination by stable emission, as stated in independent Claims 35-37. Lines 10-23 of page 9 are repeated here:

> *The reason why a slender shape is formed is to make it play a role as a spacer where, by attaching a labeled substance such as a luminescent material and the like on one end, compared with the case of attaching the labeled substance to a carrier such as a micro particle and the like, the space to the carrier and the space and distance between the labeled substances are expanded, and hence energy movement between the labeled substances and the occurrence of quenching are prevented, so that it guarantees more reliably the possibility of consistent discrimination of emissions and the like. According to the present invention, since a larger space can be obtained between the labeled substances compared with direct bonding to the carrier, interactions such as energy movement between the labeled substances, quenching (in the case of luminescent material), and the like are prevented, so that for example a number of substances, more than thousands and tens of thousands, can be discriminated consistently and with high accuracy.*

Applicants believe that support for the claim language "length of the receptor is such that a major influence by energy movement or quenching among the labeled substances does not occur, thereby enhancing discrimination by stable emission" of Claims 35-37 is present as set forth above and that the cited phrase of said claims functionally defines the length of the receptor and that the present amendment satisfies any new issue regarding written description regarding the length of the target receptor. One of ordinary skill in the art, in light of the present disclosure, would be able to determine without undue experimentation the length of a receptor "such that a major influence by energy movement or quenching among the labeled substances does not occur, thereby enhancing discrimination by stable emission" and in light of the length of the target receptor as up to 10 microns.

Since the independent claims are supported by the specification as filed, claims dependent thereon also have support. Applicants respectfully request that the rejection of Claims 8-11 and 35-37 under 35 U.S.C. §112, first paragraph, be withdrawn for the reasons cited herein.

C. Conclusion

It is believed that all matters set forth in the Final Office Action, the Advisory Action, the Communication, and the Second Advisory Action have been addressed. Further reconsideration and an early indication of the allowability of the pending claims are respectfully requested. Should the Examiner believe that an interview with Applicant's undersigned agent would expedite consideration of the pending claims, the Examiner is invited to call the undersigned agent at 512.867.8528.

Respectfully submitted,

Gloria L. Norberg

Dated: April 28, 2004

Gloria L. Norberg
Registration No. 36,706
HAYNES AND BOONE, LLP
901 Main Street - Suite 3100
Dallas, Texas 75202-3789
Telephone: (512) 867-8400
Facsimile: (214) 200-0853
ipdocketing@haynesboone.com

Research Narratives
Los Alamos National Laboratory
Center for Human Genome Studies

Center for Human Genome Studies
Los Alamos National Laboratory
P.O. Box 1663
Los Alamos, NM 87545

Larry L. Deaven
Acting Director
505/667-3912, Fax: -2891
ldeaven@telomere.lanl.gov

Lynn Clark
Technical Coordinator
505/667-9376, Fax: -2891
clark@telomere.lanl.gov

Robert K. Moyzis
Director, 198997*

*Now at University of California, Irvine

In lieu of individual abstracts, research projects and investigators at the LANL Center for Human Genome Studies are represented this narrative. More information can be found on the center's Web site.

Update

In 1997 Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, and Los Alamos National Laboratory began collaborating in a Joint Genome Institute to

Biological research was initiated at Los Alamos National Laboratory (LANL) in the 1940s, when the laboratory began to investigate the physiological and genetic consequences of radiation exposure. Eventual establishment of the national genetic sequence databank called GenBank, the National Flow Cytometry Resource, numerous related individual research projects, and fulfillment of a key role in the National Laboratory Gene Library Project all contributed to LANL's selection as the site for the Center for Human Genome Studies in 1988.

Center Organization and Activities

The LANL genome center is organized into four broad areas of research and support: Physical Mapping, DNA Sequencing, Technology Development, and Biological Interfaces. Each area consists of a variety of projects, and work is distributed among five LANL Divisions (Life Sciences; Theoretical; Computing, Information, and Communications; Chemical Science and Technology; and Engineering Sciences and Applications). Extensive interdisciplinary interactions are encouraged.

Physical Mapping

The construction of chromosome- and region-specific cosmid, bacterial artificial chromosome (BAC), and yeast artificial chromosome (YAC) recombinant DNA libraries is a primary focus of physical mapping activities at LANL. Specific work includes the construction of high-resolution maps of human chromosomes 5 and 16 and associated informatics and gene discovery tasks.

Accomplishments

- Completion of an integrated physical map (see below, left) of human chromosome16 consisting of both a low-resolution YAC contig map and a high-resolution cosmid contig map. With sequence tagged site (STS) markers provided on average every 125,000 bases, the YAC-STS map provides almost-complete coverage of the chromosome's euchromatic arms. All available loci continue to be incorporated into the map.
- Construction of a low-resolution STS map of human

implement high-throughput sequencing [see *Human Genome News* 8(2), 12].

chromosome 5 consisting of 517 STS markers regionally assigned by somatic-cell hybrid approaches. Around 95% mega-YACSTS coverage (50million bases) of 5p has been achieved. Additionally, about 40million bases of 5q mega-YACSTS coverage have been obtained collaboratively.

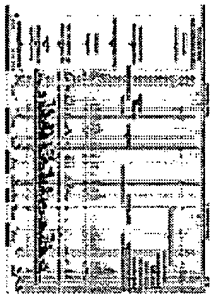
- Refinement of BAC cloning procedures for future production of chromosome-specific libraries. Successful partial digestion and cloning of microgram quantities of chromosomal DNA embedded in agarose plugs. Efforts continue to increase the average insert size to about 100,000 bases.

DNA Sequencing

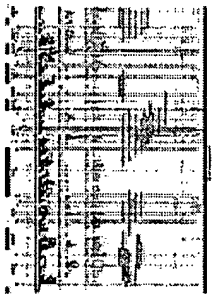
DNA sequencing at the LANL center focuses on low-pass sample sequencing (SASE) of large genomic regions. SASE data is deposited in publicly available databases to allow for wide distribution. Finished sequencing is prioritized from initial SASE analysis and pursued by parallel primer walking. Informatics development includes data tracking, gene-discovery integration with the Sequence Comparison ANalysis (SCAN) program, and functional genomics interaction.

Accomplishments

- SASE sequencing of 1.5 million bases from the p13 region of human chromosome 16.
- Discovery of more than 100 genes in SASE sequences.
- Generation of finished sequence for a 240,000-base telomeric region of human chromosome7q. From initial sequences generated by SASE, oligonucleotides were synthesized and used for primer walking directly from cosmids comprising the contig map. Complete sequencing was performed to determine what genes, if any, are near the 7q terminus. This intriguing region lacks significant blocks of subtelomeric repeat DNA typically present near eukaryotic telomeres.
- Complete single-pass sequencing of 2018 exon clones generated from LANL's flow-sorted human chromosome 16 cosmid library. About 950 discrete sequences were identified by sequence analysis. Nearly 800 appear to represent expressed sequences from chromosome 16.
- Development of Sequence Viewer to display ABI sequences with trace data on any computer having an Internet connection and a Netscape World Wide Web browser.
- Sequencing and analysis of a novel pericentromeric duplication of a gene-rich cluster between 16p11.1 and Xq28 (in collaboration with Baylor College of Medicine).



Chromosome 16 Physical Map Part 1 (80k JPG)



Technology Development

Technology development encompasses a variety of activities, both short and long term, including novel vectors for library construction and physical mapping; automation and robotics

Chromosome 16
Physical Map
Part 2 (79k JPG)

tools for physical mapping and sequencing; novel approaches to DNA sequencing involving single-molecule detection; and novel approaches to informatics tools for gene identification.

Accomplishments

- Development of SCAN program for large-scale sequence analysis and annotation, including a translator converting SCAN data to GIO format for submission to Genome Sequence DataBase.
- Application of flow-cytometric approach to DNA sizing of P1 artificial chromosome (PAC) clones. Less than one picogram of linear or supercoiled DNA is analyzed in under 3 minutes. Sizing range has been extended down to 287 base pairs. Efforts continue to extend the upper limit beyond 167,000 bases.
- Characterization of the detection of single, fluorescently tagged nucleotides cleaved from multiple DNA fragments suspended in the flow stream of a flow cytometer (see picture at left). The cleavage rate for ExoIII at 37°C was measured to be about 5 base pairs per second per M13 DNA fragment. To achieve a single-color sequencing demonstration, either the background burst rate (currently about 5 bursts per second) must be reduced or the exonuclease cleavage rate must be increased significantly. Techniques to achieve both are being explored.
- Construction of a simple and compact apparatus, based on a diode-pumped Nd:YAG laser, for routine DNA fragment sizing.
- Development of a new approach to detect coding sequences in DNA. This complete spectral analysis of coding and noncoding sequences is as sensitive in its first implementations as the best existing techniques.
- Use of phylogenetic relationships to generate new profiles of amino acid usage in conserved domains. The profiles are particularly useful for classification of distantly related sequences.

Biological Interfaces

The Biological Interfaces effort targets genes and chromosome regions associated with DNA damage and repair, mitotic stability, and chromosome structure and function as primary subjects for physical mapping and sequencing. Specific disease-associated genes on human chromosome 5 (e.g., Cri-du-Chat syndrome) and on 16 (e.g., Batten's disease and Fanconi anemia) are the subjects of collaborative biological projects.

Accomplishments

- Identification of two human 7q exons having 99% homology to the cDNA of a known human gene, vasoactive intestinal peptide receptor 2A. Preliminary data suggests that the *VIPR*2A

Flow cytometer
(18kb JPG)

gene is expressed.

- Identification of numerous expressed sequence tags (ESTs) localized to the 7q region. Since three of the ESTs contain at least two regions with high confidence of homology (~90%), genes in addition to *VIPR*2A may exist in the terminal region of 7q.
- Generation of high-resolution cosmid coverage on human chromosome 5p for the larynx and critical regions identified with Cri-du-Chat syndrome, the most common human terminal-deletion syndrome (in collaboration with Thomas Jefferson University).
- Refinement of the Wolf-Hirschhorn syndrome (WHS) critical region on human chromosome 4p. Using the SCAN program to identify genes likely to contribute to WHS, the project serves as a model for defining the interaction between genomic sequencing and clinical research.
- Collaborative construction of contigs for human chromosome16, including 1.05 million bases in cosmids through the familial Mediterranean fever (FMF) gene region (with members of the FMF Consortium) and 700,000 bases in P1 clones encompassing the polycystic kidney disease gene (with Integrated Genetics, Inc.).
- Collaborative identification and determination of the complete genomic structure of the Batten's disease gene (with members of the BDG Consortium), the gamma subunit of the human amiloride-sensitive epithelial channel (Liddle's syndrome, with University of Iowa), and the polycystic kidney disease gene (with Integrated Genetics).
- Participation in an international collaborative research consortium that successfully identified the gene responsible for Fanconi anemia type A.

Patents, Licenses, and CRADAs

- Rhett L. Affleck, James N. Demas, Peter M. Goodwin, Jay A. Schecker, Ming Wu, and Richard A. Keller, "Reduction of Diffusional Defocusing in Hydrodynamically Focused Flows by Complexing with a High Molecular Weight Adduct," United States Patent, filed December 1996.
- R.L. Affleck, W.P. Ambrose, J.D. Demas, P.M. Goodwin, M.E. Johnson, R.A. Keller, J.T. Petty, J.A. Schecker, and M.Wu, "Photobleaching to Reduce or Eliminate Luminescent Impurities for Ultrasensitive Luminescence Analysis," United States Patent, S-87, 208, accepted September 1997.
- J.H. Jett, M.L. Hammond, R.A. Keller, B.L. Marrone, and J.C. Martin, "DNA Fragment Sizing and Sorting by Laser-Induced Fluorescence," United States Patent, S.N.75,001, allowed May 1996.
- James H. Jett, "Method for Rapid Base Sequencing in DNA and RNA with Three Base Labeling," in preparation.

- Development license and exclusive license to LANL's DNA sizing patent obtained by Molecular Technology, Inc., for commercialization of single-molecule detection capability to DNA sizing.

Future Plans

LANL has joined a collaboration with California Institute of Technology and The Institute for Genomic Research to construct a BAC map of the *p* arm of human chromosome 16 and to complete the sequence of a 20-millionbase region of this map. In its evolving role as part of the new DOE Joint Genome Institute, LANL will continue scaleup activities focused on high-throughput DNA sequencing. Initial targets include genes and DNA regions associated with chromosome structure and function, syntenic break-points, and relevant disease-gene loci.

Understanding Our

Genetic Inheritance

(45k JPG)

A joint DNA sequencing center was established recently by LANL at the University of New Mexico. This facility is responsible for determining the DNA sequence of clones constructed at LANL, then returning the data to LANL for analysis and archiving.

◄ ►

◄Return to *Human Genome Project Information*
◄Return to *HGP Research Home*

Research Narratives
Lawrence Livermore National Laboratory
Human Genome Center

Human Genome Center
Lawrence Livermore National Laboratory
Biology and Biotechnology
Research Program
7000 East Avenue, L452
Livermore, CA 94551

Anthony V. Carrano
Director
510/422-5698, Fax: /423-3110
carrano1@llnl.gov

Linda Ashworth
Assistant to Center Director
510/422-5665, Fax: -2282
ashworth1@llnl.gov

In lieu of individual abstracts, research projects and investigators at LLNL Human Genome Center are represented in this narrative. More information can be found on the center's Web site.

Update:
In 1997 Lawrence Berkeley National Laboratory, Lawrence Livermore National Laboratory, and Los Alamos National Laboratory began collaborating in a Joint Genome Institute to implement high-throughput sequencing [see *Human Genome*

The Human Genome Center at Lawrence Livermore National Laboratory (LLNL) was established by DOE in 1991. The center operates as a multidisciplinary team whose broad goal is understanding human genetic material. It brings together chemists, biologists, molecular biologists, physicists, mathematicians, computer scientists, and engineers in an interactive research environment focused on mapping, DNA sequencing, and characterizing the human genome.

Goals and Priorities

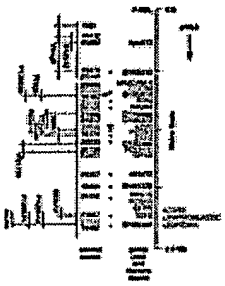
In the past 2 years, the center's goals have undergone an exciting evolution. This change is the result of several factors, both intrinsic and extrinsic to the Human Genome Project. They include: (1) successful completion of the center's first-phase goal, namely a high-resolution, sequence-ready map of human chromosome 19; (2) advances in DNA sequencing that allow accelerated scaleup of this operation; and (3) development of a strategic plan for LLNL's Biology and Biotechnology Research Program that will integrate the center's resources and strengths in genomics with programs in structural biology, individual susceptibility, medical biotechnology, and microbial biotechnology.

The primary goal of LLNL's Human Genome Center is to characterize the mammalian genome at optimal resolution and to provide information and material resources to other in-house or collaborative projects that allow exploitation of genomic biology in a synergistic manner. DNA sequence information provides the biological driver for the center's priorities:

- Generation of highly accurate sequence for chromosome 19.
- Generation of highly accurate sequence for genomic regions of high biological interest to the mission of the DOE Office of Biological and Environmental Research (e.g., genes involved in DNA repair, replication, recombination, xenobiotic metabolism, and cell-cycle control).
- Isolation and sequence of the full insert of cDNA clones associated with genomic regions being sequenced.

News 8(2), 12].

- Sequence of selected corresponding regions of the mouse genome in parallel with the human.
- Annotation and position of the sequenced clones with physical landmarks such as linkage markers and sequence tagged sites (STSs).
- Generation of mapped chromosome 19 and other genomic clones [cosmids, bacterial artificial chromosomes (BACs), and P1artificial chromosomes (PACs)] for collaborating groups.
- Sharing of technology with other groups to minimize duplication of effort.
- Support of downstream biology projects, for example, structural biology, functional studies, human variation, transgenics, medical biotechnology, and microbial biotechnology with knowhow, technology, and material resources.



Chromosome 19 Map

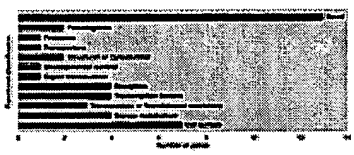
(34k JPG)

Center Organization and Activities

Completion and publication of the metric physical map of human chromosome 19 (pictured at left) in 1995 has led to consolidation of many functions associated with physical mapping, with increased emphasis on DNA sequencing. The center is organized into five broad areas of research and support: sequencing, resources, functional genomics, informatics and analytical genomics, and instrumentation. Each area consists of multiple projects, and extensive interaction occurs both within and among projects.

Sequencing

The sequencing group is divided into several subprojects. The core team is responsible for the construction of sequence libraries, sequencing reactions, and data collection for all templates in the random phase of sequencing. The finishing team works with data produced by the core team to produce highly redundant, highly accurate "finish" sequence on targets of interest. Finally, a team of researchers focuses specifically on development, testing, and implementation of new protocols for the entire group, with an emphasis on improving the efficiency and cost basis of the sequencing operation.

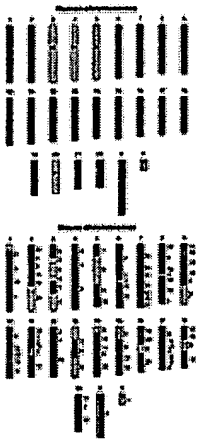


Putative-Gene Classification
(37k JPG)

Resources

The resources group provides mapped clonal resources to the sequencing teams. This group performs physical mapping as needed for the DNA sequencing group by using fingerprinting, restriction mapping, fluorescence in situ hybridization, and other techniques. A small mapping effort is under way to identify, isolate, and characterize BAC clones (from anywhere in the human genome) that relate to susceptibility genes, for example, DNA repair. These clones will be characterized and provided for sequencing and at the same time contribute to understanding the biology of the chromosome, the

genome, and susceptibility factors. The mapping team also collaborates with others using the chromosome 19 map as a resource for gene hunting.



Mouse-Human Similarities (33k

Functional Genomics

The functional genomics team is responsible for assembling and characterizing clones for the Integrated Molecular Analysis of Gene Expression (called IMAGE) Consortium and cDNA sequencing, as well as for work on gene expression and comparative genomics. The effort emphasizes genes involved in DNA repair and links strongly to LLNL's gene-expression and structural biology efforts. In addition, this team is working closely with Oak Ridge National Laboratory (ORNL) to develop a comparative map and the sequence data for mouse regions syntenic to human chromosome 19.

Informatics and Analytical Genomics

The informatics and analytical genomics group provides computer science support to biologists. The sequencing informatics team works directly with the DNA sequencing group to facilitate and automate sample handing, data acquisition and storage, and DNA sequence analysis and annotation. The analytical genomics team provides statistical and advanced algorithmic expertise. Tasks include development of modelbased methods for data capture, signal processing, and feature extraction for DNA sequence and fingerprinting data and analysis of the effectiveness of newly proposed methods for sequencing and mapping.

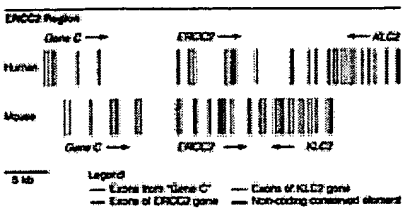
Researching Human-Mouse Homologies (20k JPG)

Instrumentation

The instrumentation group also has multiple components. Group members provide expertise in instrumentation and automation in highthroughput electrophoresis, preparation of high-density replicate DNA and colony filters, fluorescence labeling technologies, and automated sample handling for DNA sequencing. To facilitate seamless integration of new technologies into production use, this group is coupled tightly to the biologist user groups and the informatics group.

Collaborations

The center interacts extensively with other efforts within the LLNL Biology and Biotechnology Research Program and with other programs at LLNL, the academic community, other research institutes, and industry. More than 250 collaborations range from simple probe and clone sharing to detailed gene family studies. The following list reflects some major collaborations.



Comparative Sequencing (19k JPG)

- Integration of the genetic map of human chromosome 19 with corresponding mouse chromosomes (ORNL).
- Miniaturized polymerase chain reaction instrumentation (LLNL).
- Sequencing of IMAGE Consortium cDNA clones (Washington University, St.Louis).
- Mapping and sequencing of a gene associated with Finnish congenital nephrotic syndrome (University of Oulu, Finland).

Accomplishments

The LLNL Human Genome Center has excelled in several areas, including comparative genomic sequencing of DNA repair genes in human and rodent species, construction of a metric physical map of human chromosome 19, and development and application of new biochemical and mathematical approaches for constructing ordered clone maps. These and other major accomplishments are highlighted below.

- Completion of highly accurate sequencing totaling 1.6 million bases of DNA, including regions spanning human DNA repair genes, the candidate region for a congenital kidney disease gene, and other regions of biological interest on chromosome 19.
- Completion of comparative sequence analysis of 107,500 bases of genomic DNA encompassing the human DNA repair gene *ERCC*2 and the corresponding regions in mouse and hamster. In addition to *ERCC*2, analysis revealed the presence of two previously undescribed genes in all three species. One of these genes is a new member of the kinesin motor protein family. These proteins play a wide variety of roles in the cell, including movement of chromosomes before cell division.
- Complete sequencing of human genomic regions containing two additional DNA repair genes. One of these, *XRCC*3, maps to human chromosome 14 and encodes a protein that may be required for chromosome stability. Analysis of the genomic sequence identified another kinesin motor protein gene physically linked to *XRCC*3. The second human repair gene, *HHR*23A, maps to 19p13.2. Sequence analysis of 110,000 bases containing *HHR*23A identified six other genes, five of which are new genes with similarity to proteins from mouse, human, yeast, and *Caenorhabditis elegans*.
- Complete sequencing of full-length cDNAs for three new DNA repair genes (*XRCC*2, *XRCC*3, and *XRCC*9) in collaboration with the LLNL DNA repair group.
- Generation of a metric physical map of chromosome 19 spanning at least 95% of the chromosome. This unique map

incorporates a metric scale to estimate the distance between genes or other markers of interest to the genetics community.

- Assembly of nearly 45 million bases of *Eco*RI restriction-mapped cosmid contigs for human chromosome 19 using a combination of fingerprinting and cosmid walking. Small gaps in cosmid continuity have been spanned by BAC, PAC, and P1 clones, which are then integrated into the restriction maps. The high depth of coverage of these maps (average redundancy, 4.3-fold) permits selection of a minimum overlapping set of clones for DNA sequencing.
- Placement of more than 400 genes, genetic markers, and other loci on the chromosome 19 cosmid map. Also, 165 new STSs associated with pre-mapped cosmid contigs were generated and added to the physical map.
- Collaborations to identify the gene (*COMP*) responsible for two allelic genetic diseases, pseudoachondro-plasia and multiple epiphyseal dysplasia, and the identification of specific mutations causing each condition.
- Through sequence analysis of the 2A subfamily of the human cytochrome P450 enzymes, identification of a new variant that exists in 10% to 20% of individuals and results in reduced ability to metabolize nicotine and the antiblood-clotting drug Coumadin.
- Location of a zinc finger gene that encodes a transcription factor regulating blood-cell development adjacent to telomere repeat sequences, possibly the gene nearest one end of chromosome 19.
- Completion of the genomic and cDNA sequence of the gene for the human Rieske FeS protein involved in mitochondrial respiration.
- Expansion of the mouse-human comparative genomics collaboration with ORNL to include study of new groups of clustered transcription factors found on human chromosome 19q and as syntenic homologs on mouse chromosome 7.
- Numerous collaborations (in particular, with Washington University and Merck) continuing to expand the LLNL-based IMAGE Consortium, an effort to characterize the transcribed human genome. The IMAGE clone collection is now the largest public collection of sequenced cDNA clones, with more than one million arrayed clones, 800,000 sequences in public databases, and 10,000 mapped cDNAs.
- Development and deployment of a comprehensive system to handle sample tracking needs of production DNA sequencing. The system combines databases and graphical interfaces running on both Mac and Sun platforms and scales easily to handle largescale production sequencing.
- Expansion of the LLNL genome center's World Wide Web

site to include tables that link to each gene being sequenced, to the quality scores and assembled bases collected each night during the sequencing process, and to the submitted GenBank sequence when a clone is completed.

- Implementation of a new database to support sequencing and mapping work on multiple chromosomes and species. Web-based automated tools were developed to facilitate construction of this database, the loading of over 100 million bytes of chromosome 19 data from the existing LLNL database, and automated generation of Webbased input interfaces.
- Significant enhancement of the LLNL Genome Graphical Database Browser software to display and link information obtained at a subcosmid resolution from both restriction map hybridization and sequence feature data. Features, such as genes linked to diseases, allow tracking to fragments as small as 500 base pairs of DNA.
- Development of advanced micro-fabrication technologies to produce electrophoresis microchannels in large glass substrates for use in DNA sequencing.
- Installation of a new filter-spotting robot that routinely produces 6 X 6 X 384 filters. A 16 X 16 X 384 pattern has been achieved.
- Upgrade of the Lawrence Berkeley National Laboratory colony picker using a second computer so that imaging and picking can occur simultaneously.

Future Plans

Genomic sequencing currently is the dominant function of Livermore's Human Genome Center. The physical mapping effort will ensure an ample supply of sequence-ready clones. For sequencing targets on chromosome 19, this includes ensuring that the most stable clones (cosmids, BACs, and PACs) are available for sequencing and that regions with such known physical landmarks as STSs and expressed sequenced tags (ESTs) are annotated to facilitate sequence assembly and analysis. The following targets are emphasized for DNA sequencing:

- Regions of high gene density, including regions containing gene families.
- Chromosome 19, of which at least 42 million bases are sequence ready.
- Selected BAC and PAC clones representing regions of about 0.2 million to 1 million bases throughout the human genome; clones would be selected based on such high-priority biological targets as genes involved in DNA

repair, replication, recombination, xenobiotic metabolism, cell-cycle checkpoints, or other specific targets of interest.
- Selected BAC and PAC clones from mouse regions syntenic with the genes indicated above.
- Full-insert cDNAs corresponding to the genomic DNA being sequenced.

The informatics team is continuing to deploy broader-based supporting databases for both mapping and sequencing. Where appropriate, Web and Java-based tools are being developed to enable biologists to interact with data. Recent reorganization within this group enables better direct support to the sequencing group, including evaluating and interfacing sequence-assembly algorithms and analysis tools, data and process tracking, and other informatics functions that will streamline the sequencing process.

The instrumentation effort has three major thrusts: (1) continued development or implementation of laboratory automation to support highthroughput sequencing; (2) development of the next-generation DNA sequencer; and (3) development of robotics to support highdensity BAC clone screening. The last two goals warrant further explanation.

The new DNA sequencer being developed under a grant from the National Institutes of Health, with minor support through the DOE genome center, is designed to run 384 lanes simultaneously with a low-viscosity sieving medium. The entire system would be loaded automatically, run, and set up for the next run at 3-hour intervals. If successful, it should provide a 20- to 40-fold increase in throughput over existing machines.

An LLNL-designed high-precision spotting robot, which should allow a density of 98,304 spots in 96 cm^2, is now operating. The goal of this effort is to create highdensity filters representing a 10X BAC coverage of both human and mouse genomes (30,000 clones = 1X coverage). Thus each filter would provide ~3X coverage, and eight such filters would provide the desired coverage for both genomes. The filters would be hybridized with amplicons from individual or regionspecific cDNAs and ESTs; given the density of the BAC libraries, clones that hybridize should represent a binned set of BACs for a region of interest. These BACs could be the initial substrate for a BAC sequencing strategy. Performing hybridizations in parallel in mouse and human DNA facilitates the development of the mouse map (with ORNL involvement), and sequencing BACs from both species identifies evolutionarily conserved and, perhaps, regulatory regions.

Information generated by sequencing human and mouse DNA in parallel is expected to expand LLNL efforts in functional genomics. Comparative sequence data will be used to develop a high-resolution synteny map of conserved mouse-human domains and incorporate automated northern expression analysis of newly identified genes. Long range, the center hopes to take advantage of a variety of forms of expression analysis, including site-directed mutation analysis in the mouse.

Summary

The Livermore Human Genome Center has undergone a dramatic shift in emphasis toward commitment to largescale, highaccuracy sequencing of chromosome 19, other chromosomes, and targeted genomic regions in the human and mouse. The center also is committed to exploiting sequence information for functional genomics studies and for other programs, both in house and collaboratively.

◄ ►

◄Return to *Human Genome Project Information*
◄Return to *HGP Research Home*